

# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: SHIBAYAMA, SHIRO  
HIRANO, ATSUSHI  
OHNO, HIROYUKI
- (ii) TITLE OF INVENTION: A NOVEL POLYPEPTIDE AND DNAS ENCODING IT
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
  - (B) STREET: 2100 Pennsylvania Avenue, N.W.
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 20037-3202
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: JP 123155/1994
  - (B) FILING DATE: 12-MAY-1994
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (202)293-7060
  - (B) TELEFAX: (202)293-7860
  - (C) TELEX: 6491103

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 205 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Gly Thr Ala Gly Ala Met Gln Leu Cys Trp Val Ile Leu Gly Phe  
 -24 -20 -15 -10



AGAGACGGAG GCACAAGCAG AGACACATTT CAAACTGTTC CCCCCAATTC AACCACCATG	300
AGCCTGAGCA TGAGGGAAGA TGCGACCATC CTGCCCAGCC CCACGTCAGA GACTGTGCTC	360
ACTGTGGGTG CATTTGGTGT TATCAGCTTC ATTGTCATCC TGGTGGTTGT GGTGATCATC	420
CTAGTTGGTG TGGTCAGCCT GAGGTTCAAG TGTCGGAAGA GCAAGGAGTC TGAAGATCCC	480
CAGAAACCTG GGAGTTCAGG GCTGTCTGAA AGTGCTCCA CAGCCAATGG AGAGAAAGAC	540
AGCATCACCC TTATCTCCAT GAAGAACATC AACATGAATA ATGGCAAACA AAGTCTCTCA	600
GCAGAGAAGG TTCTTTAA	618

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 983 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCTGCCCGC CACATACCCA GCTGACATGG GCACCGCAGG AGCCATGCAG CTGTGCTGGG	60
TGATCCTGGG CTTCCTCTCG TTCCGAGGCC ACAACTCCCA GCCCACAATG ACCCAGACCT	120
CTAGCTCTCA GGGAGGCCCT GGCGGTCTAA GTCTGACCAC AGAGCCAGTT TCTTCCAACC	180
CAGGATACAT CCCTTCCTCA GAGGCTAACA GGCCAAGCCA TCTGTCCAGC ACTGGTACCC	240
CAGGCGCAGG TGTCCCCAGC AGTGGAAGAG ACGGAGGCAC AAGCAGAGAC ACATTTCAA	300
CTGTTCCCCC CAATTCAACC ACCATGAGCC TGAGCATGAG GGAAGATGCG ACCATCCTGC	360
CCAGCCCCAC GTCAGAGACT GTGCTCACTG TGGCTGCATT TGGTGTATC AGCTTCATTG	420
TCATCCTGGT GGTGTGGTG ATCATCCTAG TTGGTGTGGT CAGCCTGAGG TTCAAGTGTC	480
GGAAGAGCAA GGAGTCTGAA GATCCCCAGA AACCTGGGAG TTCAGGGCTG TCTGAAAGCT	540
GCTCCACAGC CAATGGAGAG AAAGACAGCA TCACCCTTAT CTCCATGAAG AACATCAACA	600
TGAATAATGG CAAACAAAGT CTCTCAGCAG AGAAGGTCT TTAAGGCAA CTTTGGGTCC	660
CCATGAGTCC AAGGATGATG CAGCTGCCCT GTGACTACAA GGAGGAAGAG ATGGAATTAG	720
TAGAGGCAAT GAACCACATG TAAATTATTT TATTGTTTCA TGTCTGCTTC TAGATCTAAA	780
GGACACTAGC ATTGCCCCAG ATCTGGGAGC AAGCTACCAA CAGGGGAGAC TCTTTCCTGT	840
ATGGACAGCT GCTGTGGAAA TACTGCCTGC TTCTCCACC TCCTCAGAGC CACAGGAAAG	900

AGGAGGTGAC AGAGAGAGAG CAAGGAAAGT GATGAGGTGG ATTGATACTT TCTACTTTGC 960  
 ATTAATAATTA TTTTCTAGCC TGC 983

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 983 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Homo sapiens  
 (H) CELL LINE: endothelial cell line of umbilical cord vein

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 27..644  
 (C) IDENTIFICATION METHOD: by similarity to some other pattern

- (ix) FEATURE:  
 (A) NAME/KEY: sig\_peptide  
 (B) LOCATION: 27..98  
 (C) IDENTIFICATION METHOD: by similarity with known sequence or  
 to an established consensus

- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 381..464  
 (C) IDENTIFICATION METHOD: by similarity with known sequence or  
 to an established consensus

- (ix) FEATURE:  
 (A) NAME/KEY: mat\_peptide  
 (B) LOCATION: 99..641  
 (C) IDENTIFICATION METHOD: by similarity with known sequence or  
 to an established consensus

- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 312..320  
 (C) IDENTIFICATION METHOD: by similarity with known sequence or  
 to an established consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCCTGCCCGC CACATACCCA GCTGAC ATG GGC ACC GCA GGA GCC ATG CAG CTG 53  
 Met Gly Thr Ala Gly Ala Met Gln Leu  
 -24 -20

TGC TGG GTG ATC CTG GGC TTC CTC CTG TTC CGA GGC CAC AAC TCC CAG 101  
 Cys Trp Val Ile Leu Gly Phe Leu Leu Phe Arg Gly His Asn Ser Gln  
 -15 -10 -5 1

CCC ACA ATG ACC CAG ACC TCT AGC TCT CAG GGA GGC CTT GGC GGT CTA 149

Pro	Thr	Met	Thr	Gln	Thr	Ser	Ser	Ser	Gln	Gly	Gly	Leu	Gly	Gly	Leu	
			5					10					15			
AGT	CTG	ACC	ACA	GAG	CCA	GTT	TCT	TCC	AAC	CCA	GGA	TAC	ATC	CCT	TCC	197
Ser	Leu	Thr	Thr	Glu	Pro	Val	Ser	Ser	Asn	Pro	Gly	Tyr	Ile	Pro	Ser	
		20					25				30					
TCA	GAG	GCT	AAC	AGG	CCA	AGC	CAT	CTG	TCC	AGC	ACT	GGT	ACC	CCA	GGC	245
Ser	Glu	Ala	Asn	Arg	Pro	Ser	His	Leu	Ser	Ser	Thr	Gly	Thr	Pro	Gly	
		35				40					45					
GCA	GGT	GTC	CCC	AGC	AGT	GGA	AGA	GAC	GGA	GGC	ACA	AGC	AGA	GAC	ACA	293
Ala	Gly	Val	Pro	Ser	Ser	Gly	Arg	Asp	Gly	Gly	Thr	Ser	Arg	Asp	Thr	
		50			55				60						65	
TTT	CAA	ACT	GTT	CCC	CCC	AAT	TCA	ACC	ACC	ATG	AGC	CTG	AGC	ATG	AGG	341
Phe	Gln	Thr	Val	Pro	Pro	Asn	Ser	Thr	Thr	Met	Ser	Leu	Ser	Met	Arg	
				70					75					80		
GAA	GAT	GCG	ACC	ATC	CTG	CCC	AGC	CCC	ACG	TCA	GAG	ACT	GTG	CTC	ACT	389
Glu	Asp	Ala	Thr	Ile	Leu	Pro	Ser	Pro	Thr	Ser	Glu	Thr	Val	Leu	Thr	
			85					90					95			
GTG	GCT	GCA	TTT	GGT	GTT	ATC	AGC	TTC	ATT	GTC	ATC	CTG	GTG	GTT	GTG	437
Val	Ala	Ala	Phe	Gly	Val	Ile	Ser	Phe	Ile	Val	Ile	Leu	Val	Val	Val	
		100					105					110				
GTG	ATC	ATC	CTA	GTT	GGT	GTG	GTC	AGC	CTG	AGG	TTC	AAG	TGT	CGG	AAG	485
Val	Ile	Ile	Leu	Val	Gly	Val	Val	Ser	Leu	Arg	Phe	Lys	Cys	Arg	Lys	
		115				120					125					
AGC	AAG	GAG	TCT	GAA	GAT	CCC	CAG	AAA	CCT	GGG	AGT	TCA	GGG	CTG	TCT	533
Ser	Lys	Glu	Ser	Glu	Asp	Pro	Gln	Lys	Pro	Gly	Ser	Ser	Gly	Leu	Ser	
		130			135					140					145	
GAA	AGC	TGC	TCC	ACA	GCC	AAT	GGA	GAG	AAA	GAC	AGC	ATC	ACC	CTT	ATC	581
Glu	Ser	Cys	Ser	Thr	Ala	Asn	Gly	Glu	Lys	Asp	Ser	Ile	Thr	Leu	Ile	
				150					155					160		
TCC	ATG	AAG	AAC	ATC	AAC	ATG	AAT	AAT	GGC	AAA	CAA	AGT	CTC	TCA	GCA	629
Ser	Met	Lys	Asn	Ile	Asn	Met	Asn	Asn	Gly	Lys	Gln	Ser	Leu	Ser	Ala	
			165					170					175			
GAG	AAG	GTT	CTT	TAA	AAGCAACTTT	GGGTCCCCAT	GAGTCCAAGG	ATGATGCAGC								684
Glu	Lys	Val	Leu	*												
		180														
TGCCCTGTGA	CTACAAGGAG	GAAGAGATGG	AATTAGTAGA	GGCAATGAAC	CACATGTAAA											744
TTATTTTATT	GTTTCATGTC	TGCTTCTAGA	TCTAAAGGAC	ACTAGCATTG	CCCCAGATCT											804
GGGAGCAAGC	TACCAACAGG	GGAGACTCTT	TCCTGTATGG	ACAGCTGCTG	TGGAAATACT											864
GCCTGCTTCT	CCCACCTCCT	CAGAGCCACA	GGAAGAGGA	GGTGACAGAG	AGAGAGCAAG											924
GAAAGTGATG	AGGTGGATTG	ATACTTTCTA	CTTTGCATTA	AAATTATTTT	CTAGCCTGC											983

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 56 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGAGCTCGTC GACAGATCTG AATTCATAT GCCCGGGGCG GCCGCACTAG TGGTAC

56

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 56 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CACTAGTGCG GCCGCCCGG GCATATGGAA TTCAGATCTG TCGACGAGCT CCTGCA

56